4400 University Drive
Fairfax, Virginia 22030

(2)

George Mason University

MODELS OF COIN-TOSSING FOR MARKOV CHAINS

by

Martin Krakowski

Report No. GMU/22474/101
December 11, 1987
(Revised)

2

<u>TECHNICAL REPORT</u>

Office of Naval Research

Contract No. N00014-86-K-0029

MODELS OF COIN-TOSSING FOR MARKOV CHAINS

by

Martin Krakowski

Report No. GMU/22474/101
December 11, 1987
(Revised)

Department of Operations Research and Applied Statistics
School of Information Technology and Engineering
George Mason University
Fairfax, VA 22030

------------------------------------ Copy No. ----------

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER	2. GOVT ACCESSION NO.
3. RECIPIENT'S CATALOG NUMBER	
4. TITLE (and Subtitle) Models of Coin-Tossing for Markov Chains	5. TYPE OF REPORT & PERIOD COVERED Technical Report
	6. PERFORMING ORG. REPORT NUMBER GMU/22474/101
7. AUTHOR(s) Martin Krakowski	8. CONTRACT OR GRANT NUMBER(s) N00014-86-K-0029
9. PERFORMING ORGANIZATION NAME AND ADDRESS Department of Operations Research and Applied Statistics George Mason University, Fairfax, Va. 22030	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS Task K-B Project 4118150
11. CONTROLLING OFFICE NAME AND ADDRESS Office of Naval Research 800 North Quincy Street Arlington, Va. 22217	12. REPORT DATE December 11, 1987
	13. NUMBER OF PAGES 45
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)	15. SECURITY CLASS. (of this report)
	15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

applied probability
computational probability
probability
stochastic modeling

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

S/N 0102-LF-014-6601

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

Abstract

Models of coin-tossing have been considered both in their own aspect or as specialized Markov chains. A common reference for both approaches is still William Feller's *Introduction to Probability Theory and its Applications*, vol. 1, 1968. A typical model calls for tossing a (biased) coin until a certain well defined stopping event, the "target," terminates the tossing; the random variable of interest, the "tally," is the number of tosses.

The stopping events in Feller and in most other references are of two sorts:

(a) A string of heads or tails of fixed size and order, e.g. hhhh, or ttttt, or ththht, etc.; or

(b) A string of r beads or of tails, which ever shows up first.

A solution is usually given as a generating function for the number of tosses; this may require tedious algebraic and numerical work in determining the distribution and moments.

In this Report we extend the Feller-type models in several ways, all believed to be new:

(1) We allow target event of variable length, e.g. h–odd # tails –h.

(2) We allow tally variables other than number of tosses, e.g. number of heads, number of tails, number of runs, number of doublets ht on way to the stopping event, etc.

(3) We allow vector-valued tallies, e.g. a vector composed of the number of heads, number of tails, number of runs, and number of triplets hht; our solution provides joint and marginal distributions, and mixed moments of the various components,

(4) A central result shows how to mechanically transform an existing solution (usually in the form of a generating function) for the number of tosses into a joint solution for the number of heads and tails. We need not even know the model to which the original solution refers. This result is of both practical and methodological interest because there are many coin-tossing problems worked out for various models in the literature, and these solutions can be transformed into joint head-and-tail solutions with minimal effort. There is an indication that such transpositions can be of value in inferential statistics.

(5) We have made a beginning towards expressing the tally process in terms of simpler building blocks, in particular geometric random variables.

(6) We show how in many instances a tossing model can be solved from the knowledge of several initial probabilities but some more work remains to be done to develop this method, which is essentially the method of indetermined coefficients.

Feller and most authors base their derivations on probability distributions, whereas we base ours on random variables and their general functionals. Probability densities and distributions, moments and generating functions are instances of such functionals.

Our notation is suitable for sums and linear combinations, including convex combinations, of random variables, and in particular for tree structures because these tend to be rich in recursions. This notation has been referred to as the omni-notation, and out method as the omni-method. The gist of the method consists in formulating the equations of mixtures and balances for an arbitrary functional of the random process of interest in place of the random process itself. This arbitrary functional can also be defined as the expectation of an arbitrary function of a random process, as we have done in other contexts including queueing theory. We refer to either definition as the omni-transform of the process.

The resulting equations are called omni-equations, and they can be specialized to distributions and generating functions, moments, costs, Pr(Z–n, modulo j) for the process or random variable z, and to other functionals. Examples of such specialization are provided as well as examples of transposing generating functions into their own omni-equivalents. The simplest and most economical way of deriving the moments of a tally variable with a known generating function is usually to first transpose this function into an omni-equation and then to find the moments algebraically or numerically.

We give examples where the omni-method applies easily whereas generating functions are cumbersome to use, e.g. the random variable "number of heads minus number of tails," which ranges over positive and negative integers, and costs which range over fractions, both positive and negative. As an indication of the power of our approach, finding the probability that the "number of heads minus number of tails" is an odd integer upon reaching the target hhh is a rather onerous combinatorial problem which reduces to several lines when using the omni-method.

Some of our results, in particular the costing applications, are applicable to more general Markov and semi-Markov chains, but this topic must be left to a future report. An attentive reader should be able, however, to make such an application by oneself.

Introduction

Various examples of coin-tossing models are discussed in Feller (1968). These models call for tossing a coin until a certain target event shows up; the random variable of interest is the number of tosses. The target events are of two kinds: (a) a string of fixed length of heads and tails in a fixed order, in particular a string of heads only or a string of tails only; and (b) a string of r heads or ρ tails, whichever shows up first. The solution is given generally in the form of a generating function for the number of tosses.

In this report we extend these models in several ways:

(1) We allow target events of variable length, e.g. h-odd #tails-h (where h stands for *heads* and t for *tails*);

(2) We allow tally variables other than the number of tosses, e.g. number of heads, number of tails, number of runs, number of doublets ht on way to targets, and others, in particular cost of tossing defined in various ways;

(3) We allow vectors of tally variables, e.g. vectors composed of number of heads, number of tails and number of runs;

(4) We show how to modify an existing solution (e.g. in forms of a generating function) for the number of heads into a bivariate solution for number of heads and number of tails; this is of both methodological and practical interest since there are in the literature various worked out solutions for the number of heads for various targets;

(5) We have made a beginning towards expressing the tally process in terms of simpler building blocks, in particular geometric random variables; and

(6) We show how, in many cases, a tossing model can be solved from the knowledge of several known initial probabilities, but more work remains to be done to extend and justify the method.

Feller works primarily with distributions, we work mainly with random variables. The notation we have introduced is suitable for sums and linear combinations, in particular convex combinations, of random variables: the omni-functions and omni-equations. However, an omni-equation can be as easily specialized to an equation in probability distributions as into an equation for moments or for generating functions. We have shown examples where the omni-method is easily applicable but generating functions are difficult to use, e.g. for "number of heads minus number of tails," and for costing variables which can be positive and negative and fractional.

This report can be of independent interest but also an introduction to Markov processes. The special structure of tossing models prompted us to prepare a separate report on this topic.

We wish to add that the method of omni-equations was originally introduced in connection with queueing problems (cf. Krakowski (July 1986, November 1986, 1987)). But the quoted references are not a prerequisite for the present report. In fact, our use of the omni-method is simpler than in queueing which involves differentiation of omni-functions.

Notation

We list here the symbols used throughout most of the report. Some symbols used only locally are for the most part omitted.

r.v. : random variable

h : toss resulting in *heads*; k*h means k successive *heads*

t : toss resulting in *tails*; k*t means k successive *tails*

gf generating function

p = $Pr(h)$, $q = Pr(t)$

N = number of tosses; $N_j = Pr(N=j)$

H = number of *heads*; H_j = number of *heads*

T = number of *tails*; N_j = number of *tails*

N_{ij} = $Pr(H=i,T=j)$

r = number of reversals h → t and t → h; the number of runs is $R=r+1$

$\psi(A)$ is an arbitrary function of A; $\psi(A,B)$ is an arbitrary function of A and B

$\delta(j)$ = 1 if $j=0$, and vanishes otherwise

$\delta(i,j)$= $\delta(i)\delta(j) = 1$ if $i=j=0$, and vanishes otherwise

E is the expectation operator

G_a is a geometric r.v. with parameter a, i.e. $Pr(G_a=j)=a(1-a)^{j-1}$

The *target* event is the event which stops the tossing *project*. Thus saying that the target is hhth means that the tossing stops as soon as we hit hhth. A target need not be of fixed length; it can be "h-odd # of t-h", or "hh or ttt whichever comes first". A tossing *model* consists of a tally, which is a random variable or a random vector of interest and a target; thus {H,T,r;hht} means that we stop tossing as soon as we hit hht and we count the number of heads and the number of tails and the number of reversals; clearly we are interested in the trivariate vector (H,T,r).

Omni–Convention An omni-equation, e.g.

$$E\psi(N) = p^2\psi(2) + pqE\psi(2+N) + qE\psi(1+N)$$

can be written as

$$\psi(N) = p^2\psi(2) + pq\psi(2+N) + q\psi(1+N).$$

This means that, following the omni-convention, we omit the operator E from the print but retain it in our mind's eye. This convention adds to the brevity of the notation. (It is analogous to Einstein's summation convention often used in tensor and matrix calculus—expectation is a summation.)

Note: $E\psi(X)$, or $\psi(X)$ in the abbreviated notation, can be thought of as an arbitrary functional of X. This simplifies, e.g., congruence equations such as (1.7) below.

Omni-Transforms and Omni-Equations

The expectation $E\psi(Z)$ of an arbitrary well-behaved function of the random variable Z will be called the *omni-transform* of Z. (For the models in this report, functions with finite expectations are well-behaved but in other contexts, say in queueing models, the existence of derivatives and their expectations is a condition of good behavior.) It is the essence of the omni-method that, along with Z, one considers all functions $\psi(Z)$.

The simplest omni-equation is

$$E\psi(A) = E\psi(B) \tag{a}$$

which says that the random variables A and B have the same distribution, be they dependent or independent. If (a) holds, if X_1 is independent of A and X_2 of B, and if $E\psi(X_1) = E\psi(X_2)$ then

$$E\psi(A+X_1) = E\psi(B+X_2) \tag{b}$$

and this is usually written as

$$E\psi(A+X) = E\psi(B+X) \tag{c}$$

Contrarywise, if (c) holds then so does (a). One has to be careful in (b) that X_1 and X_2 are independent, respectively, of A and B. Thus one must not take $X_1 = X_2 = -A$ and write (b) as

$$E\psi(A-A) = E\psi(B-A) = E\psi(0) \tag{d}$$

What we said about the simplest omni-equation (a) holds for all omni-equations.

Section 1. The Project {h}

We start with the simplest *project* in coin-tossing: toss a (p,q) coin till the target h ("heads") shows up.

The project {h} can be represented by the infinite tree of Fig. 1:

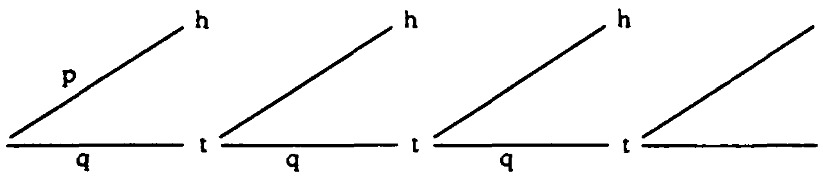
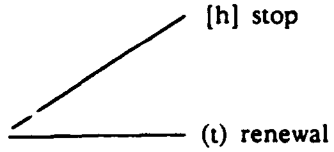


Fig. 1.1

where $p = \Pr(h)$ and $q = 1-p = \Pr(t)$ in a single toss. We assume that successive tosses are stochastically independent.

We now replace the above *infinite* tree by the *finite recursion* tree in Fig. 1.2. (Since each link pointing towards h has probability p and each link pointing towards t has probability q we can omit hte p's and q's from the diagrams.)



[ ] denotes stop

( ) denotes renewal

Fig. 1.2

If we are interested in EN we read from Fig. 1.2

$$EN = p.1 + qE(1+N^*) \tag{1.1}$$

where N^* is the remaining numer of tosses following the renewal toss (t). Although N and N^* are dependenet (in fact N^* is a subsequence of N) they are statistical replicas of each and $EN = EN^*$. Hence no confusion should result if we suppress the asterisk in (1.1) and write it as

$$EN = p.1 + qE(1+N) \qquad (1.2)$$

From (1.2) we get $\overline{N} = 1/p$.

Similarly, if we are interested in EN^2, we read from Fig. 1.2

$$EN^2 = p.1^2 + qE(1+N)^2 \qquad (1.2)$$

Similarly for EN^3 or EN^j. In fact for any *arbitrary function* $\psi(N)$ of N we read from Fig. 1.2

$$E\psi N = p\psi(1) + qE\psi(1+N) \tag{1.3}$$

an omni-equation, so-called, for the model {N; h}. The *omni-convention* improves the esthetics of the typography by removing the E but making it understood so that (1.3a) says the same thing as (1.3).

$$\psi(N) = p\psi(1) + q\psi(1+N) \tag{1.3a}$$

(Equation (1.3a)) can be also read as saying "the r.v. N is linear combination, with respective coefficients p and q, of the r.v. 1 and the r.v. $1+N^*$ where N^* is a statistical replica of N. Thus omni-equations provide a notation for sums and combinations (in particular convex mixtures).

We will refer to $E\psi(N)$ as the omni-transform of N. This transform derives its flexibility from the arbitrariness of $\psi(\ )$; and its ease of operations from the fact that expectations are simple to apply to sums and mixtures of r.v.'s.

By setting $\psi(N) = z^N$ we get from (1.3) or (1.3a)

$$EZ^N = pZ + qzEz^N$$

and

$$Ez^N = \frac{pz}{1-qz} \tag{1.4}$$

the gf of N.

By setting $\psi(N) = \delta(N-j)$, where $\psi(j) = 1$ if j=0 and $\psi(j)=0$ otherwise, we obtain

$$Pr(N=j) = p\,Pr(1=j) + q\,Pr(1+N=j) \tag{1.5}$$

since $E\delta(N-j) = Pr(N=j)$.

Denoting $N_j = Pr(N=j)$ we write (1.6) as

$$N_j = p\delta(j-1) + qN_{j-1} \tag{1.5a}$$

Of course we can read (1.5) directly from Fig. 1.2.

From (1.5a) we get successively $N_1 = p$, $N_2 = qp$, $N_3 = q^2p$ and $N_j = q^jp$.

Equation (1.5a) is recognized as the recursion of a geometric distribution with parameter p; we denote the corresponding r.v. as G_p so that, as in (1.3a)

$$\psi(G_p) = p\psi(1) + q\psi(1+G_p) \tag{1.6}$$

We can say that the N in (1.3) or (1.3a) satisfies

$$\psi(N) = \psi(G_p) \tag{1.6a}$$

By setting $\psi(N) = \sum_i \delta(N-j-ik)$ we get from (1.5) the congruence equation

$$\begin{aligned} \Pr(N \equiv j, \bmod k) = p\Pr(1 \equiv j, \bmod k) + \\ + q\Pr(1+N \equiv j, \bmod k) \end{aligned} \tag{1.7}$$

Setting k=2 and j =1 in (1.7) we have, with

$$\alpha = \Pr(N \equiv 1, \bmod 2) = \Pr(N \text{ is odd})$$

$$\alpha = P1 + q(1-\alpha)$$

and

$$\Pr(N \text{ is odd}) = \frac{1}{1+q} \tag{1.7a}$$

The question suggests itself, Are there other r.v.'s whose omni-equations are represented by the renewal tree of Fig. 1.2? Yes, there are, and this is a major reason why this paper is written. We will say that such r.v.'s are isographic with N with respect to the renewal tree of Fig. 1.2.

The Models {H; h}, {T; h}, and {H,T; h}

Let H = number of heads among the N tosses (we know that H=1 but let go on with a poker face as if we did not know it); and let T = number of tails among the N tosses. Of course, H+T = N. The omni-equations for H and for T can be read off from the tree of Fig. 1.2 which represents H and T as well as N. We thus obtain

$$\psi(H) = p\psi(1) + q\psi(H) \tag{1.8a}$$

$$\psi(T) = p\psi(0) + q\psi(1+T) \tag{1.8b}$$

In fact, the vector (H,T) is also isographic with N with respect to the tree of Fig. 2.1 so that we can combine (1.8a) and (1.8b) into a single omni-equation for (H,T) in the model {H, T; h}:

$$\psi(H,T) = p\psi(1,0) + q\psi(H,1+T) \tag{1.9}$$

We can also read off equation (1.9) directly from Fig. 1.2 rather than fuse (1.8a) and (1.8b). Both methods are instructive.

When considered separately equation (1.8a) can be simplified to

$$(1-q)\psi(H) = p\psi(1), \text{ and since } 1-q = p, \text{ to } \psi(H) = \psi(1);$$

and so $H = 1$

as it ought to be. This result is not enlightening but reassuring as to the trustiness of the method.

From (1.9) we can derive the joint distribution, the mixed moments and the joint gf for (H,T). However, the outcome is known since H = 1. Such exercises are of interest for more complex models where the results may be new.

Duration of the Toss Project. Another omni-equation which is isographic to N in Fig. 1.2 arises when we ascribe random durations to the tosses. Let

x = duration of an h-toss

y = duration of a t-toss

τ = aggregate duration of all tosses

From Fig. 1.2 we read

$$\psi(\tau) = p\psi(x) + q\psi(y+\tau), \tag{1.10}$$

a convolution equation for τ. From (1.10) we can derive successive moments of t, the Laplace transform of τ, etc. The derivation of $\bar{\tau}$ is painless. With $\psi(\tau) = \tau$ we get

$$\bar{\tau} = p\bar{x} + \frac{q}{p}\bar{y} \tag{1.11}$$

to derive the Laplace Transform for τ let $\psi(\tau) = e^{-st}$. From (1.10) we get

$$Ee^{-st} = pEe^{-sx} + qEe^{-sy}Ee^{-s}$$

$$Ee^{-st} = \frac{pEe^{-sx}}{1 - qEe^{-sy}} \tag{1.12}$$

We now ascribe a discrete cost to the tossing process in the following manner. Let x and y still denote the respective durations of an h-toss and of a t-toss. We introduce a Poisson source of intensity λ which is independent of the tossing process. Let the discrete costs be

$\#x$ = number of Poisson arrivals during x

$\#y$ = number of Poisson arrivals during y

$\#\tau$ = number of Poisson arrivals during τ

The Poisson operator "#" applied to each summand in each argument in (1.10) results in

$$\psi(\#\tau) = p\psi(\#x) + q\psi(\#y + \#\tau). \tag{1.13}$$

This is a valid equation since the Poisson operator is additive over nonoverlapping intervals, say A and B:

$$\#(A \cup B) = \#A + \#B \tag{1.14}$$

It is known that $E\#x = \lambda Ex$ and $E\#y = \lambda Ey$; and that

$$\Pr(\#x = j) = E\,\frac{e^{-\lambda x}(\lambda x)^j}{j!} \tag{1.15}$$

(cf. Gross and Harris, Section 5.1.)

From (1.13) we can find the distribution and moments of $\#\tau$. The expectation of $\#\tau$ can be obtained also from (1.11):

$$E\#\tau = \lambda E\tau = \lambda Ex + \frac{q}{p}\lambda Ey \tag{1.16}$$

Equations of the type (1.10) and (1.13) found an application in the theory of M/G/1 queues and their variants. (Cf. Krakowski July 1986; November 1986; 1987.)

A costing Model Consider the following costing model. Toss a coin till h results. The cost of tossing an h is A and the cost of tossing a t is B; A and B are random variables.

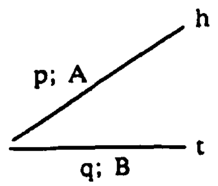


Fig. 1.3

The corresponding omni–equation is, denoting by K the cost of tossing till h,

$$\psi(K) = p\psi(A) + q\psi(B+K) \tag{1.16}$$

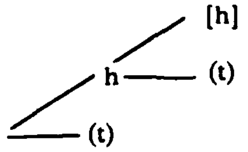
Setting $\psi(K) = K$ gives $\overline{K} = \overline{A} + \frac{q}{p}\ \overline{B}$.

Since the costs A and B can be negative so can K; hence the method of generating functions may not be applicable.

Section 2 The Project {hh}

We toss a (p,q) coin till we reach hh, i.e. two successive "heads." We start with the model {N;hh} where we tally N, the number of tosses in reaching the target pattern hh. The r.v. N has been the most commonly discussed r.v. in models of coin-tossing. Our interests, however, will not be limited to N.

The simplest renewal tree for {;hh} is



[ ] denotes stopping toss
( ) denotes renewal toss

Fig. 2.1

and its associated omni-equation is

$$\psi(N) = p^2\psi(2) + pq\psi(2+N) + q\psi(1+N) \tag{2.1}$$

(The omni-convention applies).

Note We referred to Fig. 2.1 as the simplest renewal tree. A more extended renewal tree for {N;hh} is shown in Fig. 2.1a:

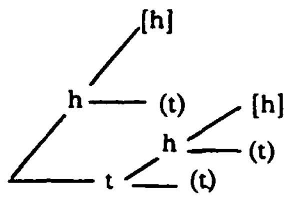


Fig. 2.1a

The omni-equation for this tree is

$$\psi(N) = p^2\psi(2) + pq\psi(2+N) + qp^2\psi(3) +$$

$$+ qpq\psi(3+N) + q^2\psi(2+N) \tag{2.1a}$$

By setting $\psi(N) = N$ we specialize (2.1) to

$$\overline{N} = 2p^2 + 2pq + pq\overline{N} + q + q\overline{N}$$

from which follows

$$\bar{N} = \frac{1}{p^2} + \frac{1}{p}\text{; for } p=q=1/2 \quad N = 6 \tag{2.2}$$

When $\psi(N) = (N-\bar{N})^2$ equation (2.1) becomes

$$E(N-\bar{N})^2 = p^2(2-\bar{N})^2 + pqE(2+N-\bar{N})^2 + qE(1+N-\bar{N})^2$$

and, since $E(N-\bar{N})=0$ we get eventually, with $\bar{N}$ from (2.2)

$$\sigma_N^2 = \frac{1}{p^4} + \frac{2}{p^3} - \frac{2}{p^2} - \frac{1}{p}\text{; for } p = q=1/2 \text{ we have } \sigma_N^2 = 22 \tag{2.3}$$

With $\psi(N) = N^k$, $k = 2, 3, 4$ etc. we can get successive moments of N from (2.1).

Letting $\psi(N) = z^N$ we get from (2.1)

$$Ez^N = p^2z^2 + pqz^2Ez^N + qzEz^N$$

from which we have the gf of N:

$$Ez^N = \frac{p^2z^2}{1-qz-pqz^2} \tag{2.4}$$

Letting $\psi(N) = \delta(N-j)$ we get, since $E\delta(N-j) = \Pr(N=j)$,

$$\Pr(N=j) = p^2\Pr(2=j) + pq\Pr(2+N=j) + q\Pr(1+N=j) \tag{2.5}$$

which, with

$$N_j \stackrel{d}{=} \Pr\,(N=j)$$

is

$$N_j = p^2\delta(j-2) + pqN_{j-2} + qN_{j-1} \tag{2.5a}$$

a recursive relation for N_j; of course $N_j = 0$ for $j < 2$.

[We can also read (2.5) directly from Fig. 2.1.]

It is easy to see directly from Fig. 2.1, or by properly specializing $\psi(N)$, that we have a relation analogeous to (2.5) in which $\Pr(N \equiv j, \bmod k)$ takes the place of $\Pr(N=j)$.

In particular, when k=2 we have $\Pr(N \equiv 1, \bmod 2) = \Pr(N \text{ is odd})$

and

$\Pr(N \equiv 0 \bmod 2) = \Pr(N \text{ is even})$

Let for brevity,

$\alpha = \Pr(N \text{ is odd}), \quad \beta = 1 - \alpha = \Pr(N \text{ is even})$

When $\psi(N) = \alpha$ we have $\psi(2+N) = \alpha$ and $\psi(1+N) = 1-\alpha$; (2.1) becomes

$$\alpha = p^2 \cdot 0 + pq \cdot \alpha + q(1-\alpha) \tag{2.6}$$

and thus

$$\alpha = \frac{q}{1+q^2} \quad \text{and} \quad \beta = 1-\alpha = \frac{1-pq}{1+q^2} \tag{2.6a}$$

Of course, we can read (2.6) directly from Fig. 2.1. When $p=q=1/2$ we have $\alpha = 2/5$ and $\beta = 3/5$.

We turn now to the processes

H = number of h-tosses upon reaching the stop (here the second h in hh)

T = number of t-tosses upon reaching the stop

Each of these processes is clearly isographic with N with respect to the renewal tree in Fig. 2.1. The omni-equations for H and for T are, as read from Fig. 2.1,

$$\psi(H) = q\psi(H) + pq\psi(1+H) + p^2\psi(2) \text{ for } \{H;hh\} \tag{2.7}$$

$$\psi(T) = q\psi(1+T) + pq\psi(1=T) + p^2\psi(0) \text{ for } \{T;hh\} \tag{2.8}$$

Since H and T are each isographic with N with respect to Fig. 2.1 then so is the vector (H,T). Hence we can fuse (2.7) and (2.8) into a bivariate omni-equation in H and T (with the omni-convention in force)

$$\psi(H,T) = q\psi(H,\ 1+T) + pq\psi(1+H,\ 1+T) + p^2\psi(2,0) \tag{2.9}$$

where $\psi(H,T)$ is an arbitrary function of H and T in the model $\{H,T;hh\}$.

When standing alone equation (2.7) can be simplified to

$$\psi(H) = p\psi(2) + q\psi(1+H) \tag{2.7a}$$

and equation (2.8) to

$$\psi(T) = p'\psi(0) + q'\psi(1+T) \text{ where } p' = p^2 \text{ and } q' = 1-p^2 \tag{2.8a}$$

Equation (2.7a) is equivalent to

$$\psi(H-1) = p\psi(1) + q\psi(H-1+1)$$

which, upon comparison with (1.6) implies that in {H; hh}

$$\psi(H-1) = \psi(G_p) \text{ and } \psi(H) = \psi(1+G_p) \tag{2.7b}$$

which says that the r.v. H–1 is a geometric r.v. with parameter p for {H;hh}. Similarly (2.8a) implies that in {T; hh}

$$\psi(T+1) = \psi(G_{p'}) \text{ where } p' = p^2 \tag{2.8b}$$

which says that the r.v. T+1 is a geometric r.v. with parameter $p' = 1-p^2$.

From (2.7b) and (2.8b) we get

$$\overline{H} = 1 + \bar{G}_p = 1 + 1/p \tag{2.10}$$

and

$$\bar{T} = -1 + \bar{G}_{p'} = -1 + 1/p' = -1 + 1/p^2 \tag{2.11}$$

and we verify that $\overline{H} + \bar{T} = 1/p + 1/p^2 = \overline{N}$ as in (2.2). From (2.7a) and (2.8a) we can find the moments and distributions of H and of T.

We turn now to the omni-equation (2.9) in order to explore the joint behavior of H and of T. Let $\psi(H,T) = HT$ and we obtain

$$\overline{HT} = q[\overline{H} + \overline{HT}\,] + pq[1 + \overline{H} + \overline{HT}\,] + p^2\cdot 0$$

from which we get, aided by (2.10) and 2.11),

$$\overline{HT} = 2/p^3 - 1/p - 1 \tag{2.12}$$

To get the joint probabilities

$$N_{ij} \triangleq \Pr(H=i,\ T=j)$$

we set $\psi(H,T) = E\delta(H-i)\delta(T-j) = Pr(H=I,T=j)$ and obtain from (2.9)

$$Pr(H=i,T=j) = qPr(H=i,1+T=j) + pqPr(1+H=i,1+T=j) + p^2Pr(i=2,\ j=0) \tag{2.13}$$

i.e.

$$N_{ij} = qN_{i,j-1}\ pqN_{i-1,j-1} + p^2\delta(i-2)\delta(j) \tag{2.13a}$$

Of course, $N_{ij} = 0$ when $i < 2$ or $j < 0$.

From (2.13a) we have $N_{2,0} = p^2$; $N_{2,1} = qN_{2,0} = p^2q$; $N_{2,j} = p^2q^j$; $N_{i,0} = 0$ for $i<2$; $N_{3,1} = q\cdot 0 + pqN_{2,0} = p^3q$; etc.

To get the bivariate gf for H and T we set

$$\psi(H,T) = z_1^H \cdot z_2^T;$$

thus (2.9) becomes

$$Ez_1^H \cdot Ez_2^T = Ez_1^H \cdot E\,z_2^T + pqz_1 \cdot z_2 \cdot Ez_1^H \cdot Ez_2^T + p_2 z_1{}^2$$

and we get the sought for generating function

$$Ez_1^H z_2^T = \frac{p^2z_1^2}{1 - qz_1 - pqz_1z_2} \tag{2.14}$$

It is often the case that the simplest way to derive a gf is to specialize an omni-equation as read from a recursion tree (or a set of such trees, as we shall see soon) or from some other structural graph. This has certainly been the case in our problems. Introduced as a prop for the span and the focus of attention, the recursion trees turned out to be of essence in the structure and classification and analysis of the models. Thus the notion of *isography* is a direct result of this deepened understanding and so is the ability to formulate multivariate problems (e.g. H and T jointly) almost as easily as univariate problems. Problems which appear weighty exercises in combinatorics (e.g., the probability that H is odd and T is even) are routinely treated as special cases of the relevant omni-equation. But more can be done still. Consider, e.g., the r.v.

$M \stackrel{d}{=}$ number of doublets ht obtained on the way to hh. Note that the M-process is isographic to N, and we get from the recursion tree in Fig. 3.1:

$$\psi(M) = q\psi(M) + pq\psi(1+M) + p^2\psi(0) \tag{2.15}$$

Equation (2.15) can be fused with (2.1) into a bivariate equation

$$\psi(N, M) = q\psi(1+N, M) + pq\psi(2+N,1+M) + p^2\psi(2,0) \tag{2.16}$$

or it can be fused with (2.9) to obtain the trivariate equation

$$\psi(H,T,M) = q\psi(H,1+T,M) + pq\psi(1+H,1+T,1+M) + p^2\psi(2,0,0) \tag{2.17}$$

Equation (2.15) when considered apart from its relation to other isographic random variables can be simplified to

$$\psi(M) = p\psi(0) + q\psi(1+M) \tag{2.15a}$$

Equation (2.15a) can be written, by adding 1 to each argument,

$\psi(1+M) = p\psi(1) + q\psi(1+1+M)$ and we recognize that

$$\psi(1+M) = \psi(G_p) \tag{2.15b}$$

which implies $\overline{M} = \bar{G}_p - 1 = 1/p - 1 = q/p$; when $p = q = 1/2$ we have $\overline{M} = 1$.

In order to show that isography should always be specified with respect to a definite renewal tree consider the model {L;hh} where

L = number of non-overlapping triplets hth on way to hh. "Non-overlapping" means, e.g., that the sequence hhthth contains only one tallied hth triplet (but the last of the six indicated tosses is an h capable of starting a new tallied hth). The tree of Fig. 1.2 is clearly not a renewal tree for the model {L; hh}. However, let us extend that tree as has been done in Fig. 1.1:

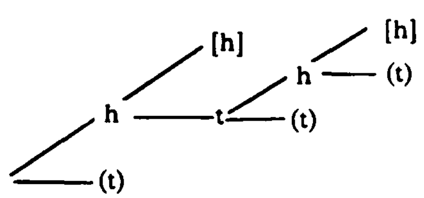


Fig. 2.2

The tree of Fig. 2.2 is clearly a renewal tree for both N and L which are therefore isographic with respect to Fig. 2.2 but not with respect to Fig. 1.2. (Note that H, T, N and L are isographic with respect to Fig. 2.2.)

Let us write down the bivariate omni–equation for (N,L) directly from Fig. 2.2. We have, and the reader should verify it,

$$\psi(N,L) = q\psi(1+N,L) + pq^2\ \psi(3+N,L) + p^2q^2\ \psi(4+N,1+L) +$$
$$p^3q\ \psi(4,1) + p^2\ \psi(2,0) \tag{2.18}$$

From (2.18) we obtain an omni–equation for N alone by disregarding the second argument in $\psi(N,L)$: $\psi(N,L) \rightarrow \psi(N)$; and we obtain an omni–equation for L by disregarding the first argument in $\psi(N.L)$: $\psi(N,L) \rightarrow \psi(L)$. The equation for L is in its nascent form (i.e. as read off from its revewal tree in Fig. 2.2).

$$\psi(L) = q\psi(L) + pq^2\psi(L) + p^2q^2\psi(1+L) + p^3q\psi(1) + p^2\psi(0) \tag{2.19}$$

which is simplified to

$$\psi(L) = \frac{q^2}{1+q}\ \psi(1+L) + \frac{pq}{1+q}\ \psi(1) + \frac{1}{1+q}\psi(0) \tag{2.20}$$

From (2.20) we can derive the moments and the distribution of L. Without deriving this distribution one can tell from the form of (2.20) that the sequence of the L_j, where $L_j = Pr(L{=}j)$, is geometrically progressing starting with L_1 and that $L_0 = 1/(1+q)$. The expected value $\bar{L}$ is

$$\bar{L} = q/(1+pq)\ ;\ \text{for } p = q = 1/2 \ \text{ we have } \bar{L} = 2/5 \tag{2.21}$$

As in Section 1, we can ascribe durations to the h–tosses and the t–tosses; and we can apply the Poisson operator to these durations but we will not pursue this matter further.

In Section 5 we will derive, in more than one way, the omni–equation for the model {N; k*h} and the model {H,T; k*h} where k*N stands for a string of k "heads".

Note What can we say about {N,L;} based on the tree of Fig. 2.1? Well, we can still write

$$\psi(L) = p^2\psi(0) + pq\psi(L_{ht}) + q\psi(L), \tag{2.22}$$

where L_{ht} is the conditional random variable denoting the number of triplets hth on way to hh *provided* that the first two tosses were h and t.

But (2.22) is one equation in two random variables and hence incomplete. It is easy to fuse N and L into $\psi(N,L)$. When standing alone (2.22) can be simplified to

$$\psi(L) = p\psi(0) + q\psi(L_{ht}) \qquad (2.22a)$$

We wish to point out, without pursuing the matter in depth, that Fig. 2.1 can be generalized to

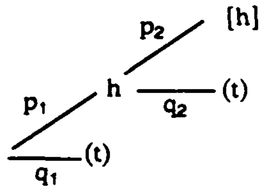


Fig. 2.3

whose omni-equation for (H,T) is

$$\psi(H,T) = q_1\ \psi(H,1+T) + p_1q_2\ \psi(1+H,1+T) + p_1p_2\ \psi(2,0) \qquad (2.23)$$

where Pr("heads") = p_2 for a toss right after an h-toss

= p_1 otherwise

and Pr("tails") = $q_2 = 1 - p_2$ for a toss right after an h-toss

$q_1 = 1 - p_1$ otherwise

We are now ready to deal with models which cannot be graphed by a single finite renewal tree. Such a model is {r; hh} where r is the number of reversals h $\to$ t or t $\to$ h; the best way to see it is to try to draw a renewal tree for this model.

Section 3 Furcation Method

Consider now the project {ht} where we toss a (p,q) coin till the couplet ht, the target, is realized. The renewal tree for this project is shown in Fig. 3.1 below:

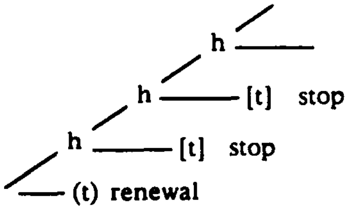


Fig. 3.1

This tree is infinite and cannot be condensed into a single finite renewal tree. (Though infinite, the above tree is still easy to handle for {N;ht} and some other models because of its simple geometry but we will not follow this possibility since it represents a special case only.)

Let us now consider the first toss as shown in Fig. 3.2:

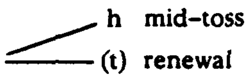
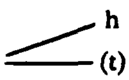


Fig. 3.2

The t-toss is a renewal, as indicated. But the h-toss is an intermediate toss, that is a toss between a renewal (or beginning) and a stop. We will refer to such a toss as a mid-toss.

As graphed in Fig. 3.2, the first toss provides us with partial information which can be stated as an omni-equation. For the model {N;ht} this equation is



$$\psi(N) = p\psi(1+N_h) + q\psi(1+N) \qquad (3.1)$$

where N_h is the remaining number of tosses provided that we already have "heads"; thus, N_h is a conditional r.v. The furcula to the left of the equation is a useful mnemonic device and we will use it repeatedly. In (3.1) we have used the fact that $\psi(N_t) = \psi(N)$ since this toss is a renewal.

Let us continue with a toss conditioned upon our having just tossed heads. Fig. 3.3 represents this conditional toss.

$$h \begin{array}{l} \nearrow h \text{ mid-toss} \\ \rightarrow [t] \text{ stop} \end{array}$$

Fig. 3.3

The omni-equation corresponding to Fig. 3.3 is

$$h \begin{array}{l} \nearrow h \\ \rightarrow t \end{array} \qquad \psi(N_h) = p\psi(1+N_h) + q\psi(1) \tag{3.2}$$

Note that from equations (3.1) and 3.2) we can reconstruct Fig. 3.1 and Fig. 3.2 (without using the furculas to the left of the equations). This goes to show that the fork equations for N and H_h retain the memory of the underlying structure. This feature is characteristic of the "furcation" method.

Equations (3.1) and (3.2) have two unknowns random variables, namely N and N_h, and we therefore expect them to be a determined set. The r.v. N_h may be thought of as an auxiliary r.v. but it may be of direct interest to some analysts; after all, the problem of dividing the pool in the midst of a gambling game arose early in the history of probability.

From Section 1 we recognize (3.2) as the equation solved by $N_h = G_q$ or $\psi(N_h) = \psi(G_q)$ where G_q is a geometric r.v. with parameter q. In fact, from the structure of the model we can reason out that

$$\psi(N) = \psi(G_p + G_q) \text{ and } \psi(N_h) = \psi(G_q) \tag{3.3}$$

The number of tosses leading to the first h is a G_p, as we well know; and the number of additional tosses leading to the first t which follows the first h is a G_q. Hence we assert (3.3). In diagram form, we have

$$\underbrace{- - - - -}_{G_p} h \underbrace{- - - - -}_{G_q} t$$

Fig. 3.4

From (3.1) and 3.2) we can obtain for N and for N_h their moments, distributions and generating functions. Let us start with $\bar{N}$ and $\bar{N}_h$. For $\psi(N) = N$ equations (3.1) and 3.2) become

$$\bar{N} = p(1+\bar{N}_h) + q(1+\bar{N}) \text{ and } \bar{N}_h = p(1+\bar{N}_h) + q\cdot 1$$

which implies

$$\bar{N}_h = 1/q \text{ and } \bar{N} = 1/p + 1/q = 1/pq \tag{3.4}$$

Of course, (3.4) also follows from (3.3).

Upon setting $\psi(N) = z^N$ equations (3.1 and (3.2) become

$$Ez^N = pzEz^{N_h} + qsEz^N \text{ and } Ez^{N_h} = pzEz^{N_h} + qz$$

which results in

$$Ez^{N_h} = \frac{qz}{1-qz} \quad \text{and} \quad Ez^N = \frac{pqz^2}{(1-qz)(1-pz)} \tag{3.5}$$

where each of the two generating functions is explicitly represented as a rational function of z.

The probabilities $(N_j) = Pr(N=j)$ and $(N_h)_j = Pr(N_h=j)$ are obtained by setting $\psi(N) = Pr(N=j)$, etc. Equations (3.1) and (3.2) become then

$$Pr(N=j) = pPr(1+N_h=j) + qPr(1+N=j) \text{ and } Pr(N_h=j) = pPr(1+N_h=j) + qPr(1=j)$$

from which all (N_j) and $(N_h)_j$ can be derived recursively starting with $(N_h)_1$.

Suppose now that we want to solve (3.1) and (3.2) for N, i.e. to find an omni-equation for N alone. From (3.1) we get

$$p\psi(1+N_h) = \psi(N) - q\psi(1+N)$$

and, dividing throughout by p and subtracting 1 from each argument, we have

$$\psi(N_h) = \frac{1}{p}\psi(N-1) - \frac{q}{p}\psi(N)$$

Substituting the last two equations into (3.2) we get an uncoupled equation for $\psi(N)$ in {N;ht}:

$$\psi(N) = \psi(1+N) - pq\psi(2+N) + pq\psi(2) \tag{3.6}$$

Setting $\psi(N) = Pr(N=j) = (N)_j$ in (3.6) we get

$$(N)_j = (N)_{j-1} - pq(N)_{j-2} + pq\delta(j-2) \quad \text{where } (N)_j = 0 \text{ if } j \leq 1 \tag{3.7}$$

From (3.7) we have $(N)_2 = pq$, $(N)_3 = pq$, $(N)_4 = pq(1-pq)$, $(N)_5 = pq(1-2pq)$, etc.

Of course, in numerical work p and q = 1-p are fixed and the recursion involves only real numbers, not symbols. Cf. (3.18b) for an explicit solution of (3.7). One can also derive (3.6) from (3.5). In fact, generating functions can be very effective in such uncoupling of the random variables entering omni-equations. This is not really surprising since generating functions have been applied to the solution of difference equations of which omni-equations are a special species.

A solution to a coin-tossing model can be given by an explicit formula for the probabilities (or moments), by recursion formulas for the probabilities, by generating functions, or by omni-equations. It is often very easy to switch from one representation to the other, especially among omni-equations, generating function, and recursions for probabilities. The Table 3.1 below shows corresponding terms for the three representations for A, a r.v.

Omni-Term	**Generating Term**	**Probability Term**
$\psi(A)$	Ez^A	$A_j \overset{d}{=} Pr(A=j)$
$\psi(1+A)$	zEz^A	A_{j-1}
$\psi(2+A)$	z^2Ez^A	A_{j-2}
$\psi(k+A)$	z^kEz^A	A_{j-k}
$\psi(0)$	1	$\delta(j) \overset{d}{=} \begin{cases} 1 & \text{if } j=0 \\ 0 & \text{if } j\neq 0 \end{cases}$
$\psi(1)$	z	$\delta(j-1)$
$\psi(2)$	z^2	$\delta(j-2)$
$\psi(k)$	z^k	$\delta(j-k)$
$\psi(G_a)$	$\dfrac{az}{1-(1-a)z}$	$(G_a)_j = a(1-a)^{j-1}$

Table 3.1

Writing (3.5) as a linear combination of terms

$$Ez^{N_h} - qzEz^{N_h} = qz \quad \text{and} \quad Ez^N - zEz^N + pqz^2Ez^N = pqz^2 \tag{3.5a}$$

we can transpose them, term by term, into their omni-form with the aid of Table 3.1:

$$\psi(N_h) - q\psi(N_h+1) = q\psi(1) \quad \text{and} \quad \psi(N) - \psi(N+1) + pq\psi(N+2) = pq\psi(2) \tag{3.8}$$

the second of which is equivalent to (3.6). Equation (3./8) or (3.5a) can likewise be transposed into its equivalent recursion in probabilities:

$$(N_h)_j - q(N_h)_{j-1} = q\delta(j-1) \quad \text{and} \quad (N)_j - (N)_{j-1} + pq(N)_{j-2} = pq\delta(j-2) \tag{3.9}$$

the second equation of (3.9) being equivalent to (3.7).

It is important that we are able to switch easily between the omni-form, the generating function and probability recursions of a model. We may want to have the omni-equation to calculate the moments or the probabilities modulo some integer but we happen to have the generating function or a probability recursion. The literature abounds in generating functions upon which we can draw. Moreover, even a skilled practitioner of omni-equations may encounter a problem where a probability recursion is easier to derive than an omni-equation. And even in essentially omni-work we may want to solve a system of fork equations, often much bigger than (3.1) and (3.2), where generating functions may be the method of choice.

We now prove a theorem both by means of omni-equations and generating functions, the latter aided by a partial fraction expansion.

Theorem Let G_a and G_b be independent geometric random variables with respective parameters a and b where $a \neq b$. Then

$$\psi(G_a + G_b) = \frac{b}{b-a}\psi(G_a) + \frac{a}{a-b}(\psi G_b) \tag{3.10}$$

Equation (1.6) states that

$$\psi(G_a) = a\psi(1) + (1-a)\psi(1+G_a) \tag{3.11a}$$

$$\psi(G_b) = b\psi(1) + (1-b)\psi(1+G_b) \tag{3.11b}$$

Shift now the first of the above equations by G_b and a the second by G_a:

$$\psi(G_a+G_b) = a\psi(1+G_b) + (1-a)\psi(1+G_a+G_b) \tag{3.12a}$$

$$\psi(G_b+G_b) = b\psi(1+G_a) + (1-b)\psi(1+G_b+G_a) \tag{3.12b}$$

By eliminating the term $\psi(1+G_a+G_b)$ from the above two equations we get

$$(a-b)\psi(G_a+G_b) = a(1-b)\psi(1+G_b) - b(1-a)\psi(1+G_a) \tag{3.13}$$

From (3.13) and (3.11a) and (3.11b) we have (3.10). Note that the right-hand side of (3.10) is a linear but not convex sum since either b–a or else a–b is negative. Thus, as has been pointed out by Botta and Harris (1986) linear but nonconvex sums of random variables arise in contexts where the original formulation of a problem is in the form of sums and convex combinations of random variables.

It can be shown by shifting (3.10) by G_c that one has

$$\psi(G_a + G_b + G_c) = \frac{bc}{(b-a)(c-a)}\,\psi(G_a) + \frac{ca}{(a-b)(c-b)}\,\psi(G_b) + \frac{ab}{(a-c)(b-c)}\,\psi(G_c) \tag{3.14}$$

where a, b, and c are distinct. (It should be added that it is possible to interpret G_a so that $\underline{a}$ need not be positive and less than 1 but we cannot delve here into this matter.)

We derive now an expression equivalent to (3.10), in terms of generating functions. Since G_a and G_b are assumed independent, and since

$$Ez^{G_a} = \frac{az}{1-(1-az)} \quad \text{and} \quad Ez^{G_b} = \frac{bz}{1-(1-bz)} \tag{3.15}$$

we have

$$Ez^{G_a+G_b} = Ez^{G_a}Ez^{G_b} = \frac{az}{1-(1-az)}\;\frac{bz}{1-(1-bz)} \tag{3.16}$$

With the aid of the partial fraction expansion and its algebra we get

$$\frac{az \cdot bz}{[1-(1-az)][1-(1-bz)]} = \frac{b}{b-a}\;\frac{az}{1-(1-a)z} + \frac{a}{a-b}\;\frac{bz}{1-(1-bz)} \tag{3.17}$$

From (3.15), (3.16) and (3.17) we have

$$Ez^{G_a+G_b} = \frac{b}{b-a}\,Ez^{G_a} + \frac{a}{a-b}\,Ez^{G_b} \tag{3.18}$$

which translates (3.10) into the language of generating functions.

Of course, (3.10) yields (3.16) upon setting $\psi(G_a) = z^{G_a}$, etc., but we wanted a purely gf derivation. Setting in (3.10) $\psi(G_a) = Pr(G_a{=}j)$ etc., we get

$$\Pr(G_a + G_b = j) = \frac{b}{b-a}\Pr(G_a = j) + \frac{a}{a-b}\Pr(G_b = j) = \tag{3.18a}$$

$$= \frac{b}{b-a}(1-a)^{j-1}a + \frac{a}{a-b}(1-b)^{j-1}b$$

Setting a = p and b = q in (3.18a) we have for the {N;ht}, in view of (3.3),

$$N_j = \frac{q}{q-p}\,q^{j-1}p + \frac{p}{p-q}\,p^{j-1}q \tag{3.18b}$$

an explicit expression for N_j.

We return now to the project {hh}. Although the model {N;hh} can be represented by a single finite renewal tree the model {r;hh} (where r = number of reversals $h \to t$ and $t \to h$ in the project {hh}) cannot be so represented. We will apply the general forking method to analyze it. The first toss is described by the omni-equation

$$\begin{smallmatrix}\nearrow h\\ \to t\end{smallmatrix} \qquad \psi(r) = p\psi(r_h) + q\psi(r_t) \tag{3.19}$$

where r_h = number of remaining reversals if the first toss is h

r_t = number of remaining reversals if the first toss is t.

The fork-equations conditioned upon having just tossed h, or else just tossed t, are

$$h\begin{smallmatrix}\nearrow [h]\\ \to t\end{smallmatrix} \qquad \psi(r_h) = p\psi(0) + q\psi(1+r_t) \tag{3.20}$$

$$t\begin{smallmatrix}\nearrow h\\ \to t\end{smallmatrix} \qquad \psi(r_t) = p\psi(1+r_h) + q\psi(r_t) \tag{3.21}$$

The three equations (3.19), (3.20) and (3.21) allow us to derive the distributions of r, or r_h and of r_t; their moments and their generating functions; and other functionals. Note that

$R = r+1$, $R_h = r_h$ and $R_t = r_t$

where R = number of runs in any sequence; R_h and R_t are the corresponding conditional random variables.

It is easy to verify that

$$\bar{r} = (2-p2)/p, \ \bar{r}_h = 2q/p \quad \text{and} \quad \bar{r}_t = (1+q)/p \tag{3.21a}$$

When $p = q = 1/2$ we get $\bar{r}_h = 2.5$, $\bar{r}_h = 2$, and $\bar{r}_t = 3$.

We derive the generating functions for r, for r_h and for r_t by setting $\psi(r) = Ez^r$, etc. After a little algebra we get

$$Ez^r = \frac{p^2+pqz}{1-qz^2}; \quad Ez^{r_h} = \frac{p}{1-qz^2}; \quad Ez^{r_t} = \frac{pz}{1-qz} \tag{3.22}$$

Aided by Table 3.1 we recycle these gf's into uncoupled omni-equations for r, for r_h and for r_t:

$$\psi(r) = p^2\psi(0) + pq\psi(1) + q\psi(r+2) \tag{3.23}$$

$$\psi(r_h) = p\psi(0) + q\psi(2+r_h) \tag{3.24}$$

$$\psi(r_t) = p\psi(1) + q\psi(2+r_t) \tag{3.25}$$

From Table 3.1, or by setting $\psi(r) = Pr(r=j) = (r)_j$ in (3.23)

we get

$$(r)_j = p^2\delta(j) + pq\delta(j-1) + q(r)_{j-2} \tag{3.26}$$

and successively

$$(r)_0=p^2, \ (r)_1=pq \ \ (r)_2=qp^2, \ \ (r)_3=pq^2; \ (r)_{2j}=q^jp^2, \ (r)_{2j+1} = pq^{2j+1}$$

We find $\alpha \overset{d}{=} Pr(r \text{ is odd})$ by setting $\psi(r) = Pr(r \text{ is odd})$ in (3.23):

$$\alpha = pq + q\alpha, \text{ hence } \alpha = q \tag{3.27}$$

A major advantage of the omni-method is its ease in formulating multivariate problems. Suppose that we are interested in jointly treating N and r in the project {hh}; we designate the model of interest as {N,r;hh}. Let us proceed in developing the elementary forks (the furculae) and their associated omni-equations. The first toss results in

$$\begin{smallmatrix} h \\ t \end{smallmatrix} \quad \psi(N,r) = p\psi(1+N_h,r_h)+q\psi(1+N_t,r_t) \tag{3.28}$$

Unlike in the model {N;hh} we now distinguish between $\psi(N_t)$ and $\psi(N)$ because of the interdependence between N_t and r_t. The fork-equations conditional upon the prior toss having been a non-stopping h, or else t, are

$$h\begin{smallmatrix} [h] \\ t \end{smallmatrix} \quad \psi(N_h, r_h) = p\psi(1,0) + q\psi(1+N_t,1+r_t) \tag{3.29}$$

$$t\begin{smallmatrix} h \\ t \end{smallmatrix} \quad \psi(N_t,r_t) = p\psi(1+N_h,1+r_h) + q\psi(1+N_t,1+r_t) \tag{3.30}$$

From the above equations we can find the joint (and marginal) distributions for N and r; for N_h and r_h; and for N_t and r_t; we can find recursively all the mixed moments for the pairs just indicated; we can find the bivariate gf's for these pairs; and we can find some other bivariate functionals.

At this point it is worth indicating how to combine N and r (and in other problems any number of random variables) into a random vector. Define

$$V = (N,r),\ V_h = (N_h,r_h) \text{ and } V_t = (N_t,r_t)$$

$$\sigma_{ij} = (i,j)$$

The fork-equations (3.28), (3.29) and (3.30) can be extended, using this vector notation, to

$$\psi(V) = p\psi(\sigma_{10}+V_h) + q\psi(\sigma_{10}+V_t) \tag{3.28a}$$

$$\psi(V_h = p\psi(\sigma_{10}) + q\psi(\sigma_{11} + V_t) \tag{3.29a}$$

$$\psi(V_t) = p\psi(\sigma_{11}) + q\psi(\sigma_{10} + V_t) \tag{3.30a}$$

The above form should be somewhat more comfortable for uncoupling the set (3.28), (3.29) and (3.30) in a way analogous to deriving (3.8) from (3.1) and (3.2) via (3.5a) for the model {N;ht}.

Many authors on probability consider that their task is completed upon deriving a generating function or a set of such functions (or the kindred Laplace transforms). We claim the same privilege at this time.

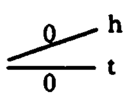
The flexibility of the omni-method and of the fork-equations enables us to replace or supplement the process N, number of tosses, by the process K, the cost of tossing. We can think of N as the cost when each toss costs exactly 1 unit. Suppose now that the cost of a toss depends on whether it is $h \rightarrow t$ or $t \rightarrow t$, etc. E.g. consider the cost-matrix of Table 3.2:

From \ To	h	t
Start	0	0
h	1	0
t	−2	1

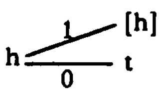
Table 3.2

More complex cost matrices can take more tossing history into account. (Such costing patterns were considered in some of our work on Markov chains, to be published in the future.)

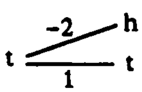
Although the model {N;hh} is described by a single finite renewal tree the model {K;hh}, with the K–matrix in Table 3.2 cannot be reduced to such a tree. We need several fork equations. It is not necessary to know in advance the number of elementary forks, these forks arise in the process of developing the equations. In our case we have for the cost K



$$\psi(K) = p\psi(K_h) + q\psi(K_t) \tag{3.31}$$



$$\psi(K_h) = p\psi(1) + q\psi(K_t) \tag{3.32}$$



$$\psi(K_t) = p\psi(-2+K_h) + q\psi(1+K_t) \tag{3.33}$$

Setting $\psi(K) = K$, etc., we get

$$\overline{K} = p\overline{K}_h + q\overline{K}_t;\ \ \overline{K}_h = p + q\overline{K}_t;\ \overline{K}_t = -2p + p\overline{K}_h + q + q\overline{K}_t$$

from which we find when $p = q = 1/2$

$$\overline{K} = -1/2;\ \overline{K}_h = 0;\ \overline{K}_t = -1$$

Since such costing models involve generally both positive and negative costs generating functions may be difficult or impossible to apply.

We formulate now the bivariate model {N,K;hh}. Its fork-equations corresponding to (3.31), (3.32) and (3.33) are

$$\psi(N,K) = p\psi(1+N_h,K_h) + q\psi(1+N_t,K_t)$$

$$\psi(N_h,K_h) = p\psi(1,1) + q\psi(1+N_t,K_t)$$

$$\psi(N_t,K_t) = p\psi(1+N_h,2+K_h) + q\psi(1+N_t,1+K_t)$$

We shall not follow this model further.

Section 4 On Converting Omni-Equations in N into Omni-Equations in H and T

It seems to have escaped general notice that the equations for N, number of tosses in a project, carry enough information to be transformed into bivariate equations in H and T, the number of h-tosses and the number of t-tosses. This is true whether the N-equations are in the omni-form, g.f. form or in distribution form. However, it is easier to observe this convertibility from N to (H,T) for omni-equations.

For didactic purposes let us start with (2.9) alias (4.1):

$$\psi(H,T) = q\psi(H,1+T) + pq\psi(1+H,1+T) + p^2\psi(2,0) \qquad (2.9) = (4.1)$$

Specializing the general function $\psi(H,T)$ to $\psi(H+T) = \psi(N)$ we have

$$\psi(N) = q\psi(1+N) + pq\psi(2+N) + p^2\psi(2) \qquad (2.1) = (4.2)$$

The question suggests itself now, How do we rederive (4.1) from (4.2) if this is at all possible? For generally, operations which mingle variables tend to obliterate information; and the passage from (4.1) to (4.2) comingles H and T by summing the two into N. But a close comparison of (4.1) and (4.2) reveals that in our case there is enough special structure in (4.2) to allow us to derive from it (4.1). Compare the first right-hand term of (4.1) with its homologue in (4.2); we have

$$q\psi(H,1+T) \quad \text{and} \quad q\psi(1+N)$$

Each of these terms has the coefficient q, a reminder that the last toss was a t-toss. The summand 1 in the argument (1+N) tells that one toss has taken place; and summand 1 in the bivariate argument (H,1+T) tells that a t-toss has taken place. Thus, the term $q\psi(H, 1+T)$ says twice that the toss was "tails" whereas the term $q\psi(1+N)$ says it only once. But one hint is enough for the wise, as the saying goes, and thus we know that

$$q\psi(1+N) \sim q\psi(H,1+T) \qquad (4.3)$$

if indeed (4.2) can be transformed into (4.1).

Let us now compare

$pq\psi(1+H,1+T)$ with $pq\psi(2+N)$.

The coefficient pq in each of the two terms bears witness to the fact that one h-toss and one t-toss took place. The argument (2+N) confirms that altogether two tosses took place; the bivariate argument $(1+H,1+T)$ specifies again that an h–toss and a t–toss took place. Therefore the term $pq\psi(2+N)$ carries enough information for its conversion:

$$pq\psi(2+N) \sim q\psi(1+H,1+T) \tag{4.4}$$

The terms $p^2\psi(2,0)$ in (4.1) and $p^2\psi(2)$ in (4.2) each says through its coefficient, that the toss couple hh took place. The argument in $p^2\psi(2)$ confirms the number of tosses, while the argument in $\psi(2,0)$ specifies that hh occurred. Thus

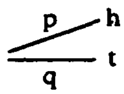
$$p^2\psi(2) \sim p^2\psi(2,0) \tag{4.5}$$

Thus we have transformed (4.2) into (4.1). A similar argument allows us to transform any linear omni–equation in N into an equation in (H,T) provided that in the N–equation each term is of the form $p^a q^b \psi(a+b+N)$ or $p^a q^b \psi(a+b)$; for such terms the correspondence with (H,T) terms is

$$p^a q^b \psi(a+b+N) \sim p^a q^b \psi(a+H,b+T) \text{ and } p^a q^b \psi(a+b) \sim p^a q^b \psi(a,b) \tag{4.6}$$

We will call such terms in N or in (H,T) *upright*. A linear omni–equation shall be called upright if each of its terms is upright. Thus, (4.2) and (4.1) are upright.

It is easy to see that each model for N can be formulated as an upright omni–equation. This is clearly seen when one starts with a renewal tree (it is even clear for an infinite tree such as Fig. 1.1) and it is clear that each nascent fork equation is also upright. For the general form of such a fork–equation is



$$\psi(A) = p\psi(1+A_h) + q\psi(1+A_t) \tag{4.7}$$

where the r.v. A is N or N_h or N_t or N_{ht} or N_{th} or another conditioned N. (If $A = N_{ht}$ then $A_h = N_{hth}$; etc.) One can also see that when uncoupling a set of such forks into a set of univariate equations (one equation for N, another for N_h, etc.) one can preserve the uprightness by careful algebra.

But what about those omni–equations for N which contain non–upright terms? For example, the omni–equation (3.6) for {N;ht}

$$\psi(N) = \psi(1+N) - pq\psi(2+N) + pq\psi(2) \qquad (3.6)= (4.8)$$

contains a term which is not upright, namely $\psi(1+N)$. Without disturbing the upright terms we multiply $\psi(1+N)$ by p+q in (4.8) thus obtaining

$$\psi(N) = p\psi(1+N) + q\psi(1+N) - pq\psi(2+N) + pq\psi(2) \tag{4.9}$$

which is an upright equation. Converting (4.9) term by term we apply the transitions

$$\begin{aligned}
\psi(N) &\rightarrow \psi(H,T)\\
p\psi(1+N) &\rightarrow p\psi(1+H,T)\\
q\psi(1+N) &\rightarrow q\psi(H,1+T)\\
-pq\psi(2+N) &\rightarrow -pq\psi(1+H,1+T)\\
pq\ (2) &\rightarrow pq\psi(1,1)
\end{aligned}$$

and obtain an upright omni-equation in H,T for the model {N;ht}

The polynomial p+q is the only polynomial in p and q which is of degree one and which equals one. Likewise, to make the term $\psi(3+N)$ upright we multiply it by $(p+q)^3$. (The omni-equations in N where each term applies to a single path are born upright; but where the terms corresponding to some different paths but with identical arguments are combined the uprightness may be lost.)

The notion of uprightness applies also to gf terms and equations. We say that a gf-term for N is upright if it has the form

$$(pz)^a Ez^N \tag{4.10}$$

and that a gf-term for H,T is upright if it has the form

$$(pz_1)^a (qz_2)^b Ez_1^H \cdot Ez_2^T. \tag{4.11}$$

It is easy to see that upright omni-terms and upright gf-terms correspond to each other; this holds for terms in N and for terms in H,T.

The following diagram illustrates the relation between upright omni-terms and gf-terms; and between N-terms and H,T-terms:

$$\begin{tikzcd}
p^a q^b \psi(a+b+N) \arrow[r, leftrightarrow] \arrow[d, leftrightarrow] & p^a q^b \psi(a+H,b+t) \arrow[d, leftrightarrow] \\
p^a q^b z^{a+b} Ez^N \arrow[r, leftrightarrow] & (pz_1)^a (qz_2)^b Ez_1^H Ez_2^T
\end{tikzcd}$$

Fig. 4.1

The above diagram remains valid when we set N = 0; H = 0; and T = 0.

[We could have defined an upright gf–term as a result of setting $\psi(N) = Ez^N$ or setting $\psi(H,T) = Ez_1^H \cdot Ez_2^T$ in upright omni–terms.]

We can likewise define uprightness for linear terms in recursive relations for probability densities by specializing $\psi(N) = Pr(N=j)$ and $Pr(H=i,T=j) = N_{ij}$ in omni–terms. Denoting as before $(N)_j = Pr(N=j)$, and $(N)_{ij} = Pr(H=i,T=j)$ we see that upright terms for probability densities are

$$p^a q^b N_{j-a-b} \quad \text{and} \quad p^a q^b N_{i-a,j-b} \; ;$$

and

$$p^a q^b \delta_{j-a-b} \quad \text{and} \quad p^a q^b \delta_{i-a}\, \delta_{j-b} \, . \qquad 4.12$$

Example from Feller Consider the model {N;k*h} where k*h stands for a string of k "heads." Feller, p.323, equation (7.6), gives two equivalent expressions for the g.f. of N:

$$Ez^N = \frac{p^k z^k}{1 - qz(1+pz ++ p^{k-1} z^{k-1})}, \qquad (4.13a)$$

and

$$Ez^N = \frac{p^k z^k \,(-pz)}{1 - z + qp^k z^{k+1}} \qquad (4.13b)$$

We notice that (4.13a) is upright, i.e. each of its terms is upright whereas (4.13b) has one term which is not upright, namely $-z$ in the denominator. Not surprisingly, (4.13b) is a result of summing the (finite) geometric series in the denominator of (4.13a) and is useful for large k.

The conversion of (4.13a) into the gf for the model {H,T;k*h} is straightforward, and we show the result only:

$$Ez_1^H \cdot z_2^T = \frac{p^r z_1{}^r}{1 - qz_2(1+pz_1 ++ p^{k-1} z_1{}^{k-1})} \qquad (4.14a)$$

In order to convert (4.13b) into the gf in H,T we first make it upright by replacing the term $-z$ in the denominator by $-(p+q)z$, a very gentle intervention:

$$Ez^N = \frac{p^k z^k - p^{k+1} z^{k+1}}{1 - pz-qz + qp^k z^{k+1}} \tag{4.13c}$$

It is important to note that in order to transpose (4.13a) and (4.13b) which are N–equations, into (4.14a) and (4.14b) below which are (H,T)–equations we need not know the target of the project, here k^*h; we do have to know that we start with N–equations.

From (4.13c) we derive the corresponding gf in H,T (Fig. 4.1 may be used for term correspondences) and obtain

$$Ez_1^H \cdot z_2^T = \frac{p^k z^k - p^{k+1} z^{k+1}}{1 - pz_1 - qz_2 + (pz_1)^k qz_2} \tag{4.14b}$$

From (4.14b) we now derive the omni–equation in H,T. We obtain, by first clearing the fraction, converting each term as indicated in Fig. 4.1

$$\begin{aligned}\psi(H,T) &= p\psi(1+H,T) + q\psi(H,1+T) - p^k q\psi(k+H,1+T) \\ &\quad + p^k \psi(k,0) - p^{k+1}\psi(k+1,0)\end{aligned} \tag{4.15}$$

Equation (4.15) can be converted into a recursion equation for $N_{ij} = Pr(H=i,T=j)$ by setting $\psi(H,T)=Pr(H=i,T=j)$; then

$$\begin{aligned}N_{ij} &= pN_{i-1,j} + qN_{i,j-1} - p^k qN_{i-k,j-1} \\ &\quad + p^k \delta(i-k)\delta(j) - p^{k+1}\delta(j-k-1)\delta(j) \quad .\end{aligned} \tag{4.16}$$

We will not delve into the intricacies of how to evaluate numerically the N_{ij} or the mixed moments of H and T. Our aim was to show how to transform an equation in N into an equation in T,H. One simple result deserves being mentioned. From (4.15) we extract the omni–equation for T alone by letting $\psi(H,T) \to \psi(T)$; after some simplification we have

$$\psi(T) = (1-p^k)\psi(1+T) + p^k \psi(0) \tag{4.17}$$

and we recognize that $\psi(T) = \psi(G_{p'} - 1)$ where $p' = p^k$; cf. (1.6).

From (4.15) we get the separate omni–equations for $\psi(H)$ and $\psi(T)$ by specializing $\psi(H,T) \to \psi(H)$ and $\psi(H,T) \to \psi(T)$:

$$\psi(H) = p\psi(1+H) + q\psi(H) - p^k q\psi(k+H) + p^k\psi(k) - p^{k+1}\psi(k+1)$$

$$\psi(T) = p\psi(T) + q\psi(1+T) - p^k q\psi(1+T) + p^k\psi(0) - p^{k+1}\psi(0)$$

which simplify to

$$\psi(H) = \psi(1+H) - p^{k-1}q\psi(k+H) + p^{k-1}\psi(k) - p^k\psi(k+1) \tag{4.15a}$$

$$\psi(T) = (1-p^k)\psi(1+T) + p^k\psi(0) \tag{4.17a}$$

Now, from either (4.13a) or (4.15) we have, since H+T=N,

$$\psi(N) = \psi(1+N) - qp^k\psi(1+k+N) + p^k\psi(k) - p^{k+1}\psi(k+\ ) \tag{4.18}$$

From (4.17a) we recognize that, with $p' = p^k$,

$$\psi(T) = \psi(G_{p'}-1);\ \text{cf. } (1.6) \tag{4.19}$$

Suppose now that we are interested in the r.v. H – T. By specializing $\psi(H,T) \to \psi(H-T)$ and denoting H – T = L we get from (4.15)

$$\psi(L) = p\psi(1+L) + q\psi(-1+L) - p^k q\psi(k-1+L) + p^k\psi(k) - p^{k+1}\psi(k+1) \tag{4.20}$$

We obtain the moments of L from (4.20) like from any other linear omni-equation. Thus we get by setting $\psi(L) = L$.

$$\bar{L} = \frac{p-q}{q}\left[\frac{1}{p^k} - 1\right] \tag{4.21}$$

There is no gf of L corresponding to (4.20) since L is not non-negative. A result like (4.21) would have to be derived from the bivariate gf for H and T, i.e. (4.14b), or from the joint distribution of H and T. The recursion equation for the distribution of L is obtained from (4.20) by setting $\psi(L) = \Pr(L=j) = (L)_j$:

$$L_j = p(L)_{j-1} + q(L)_{j+1} - p^k q(L)_{j+1-k} + p^k\delta(j-k) - p^{k+1}\delta(j-1-k) \tag{4.22}$$

It can be shown that for a given k one can, starting with k–1 consecutive values of known $(L)_j$ (e.g. $(L)_0$, $(L)_1$ ...$(L)_{k-1}$), derive the other values. And it is easy to see that for k=1 we have L=l; and that for k=2 we have L=2; and that for k=3 there is neither a lower nor an upper bound on $(L)_j$.

In order to complete honorably the Feller example we derive for N its omni–equation and its distribution. From Fig. 4.2 below we read (4.23)

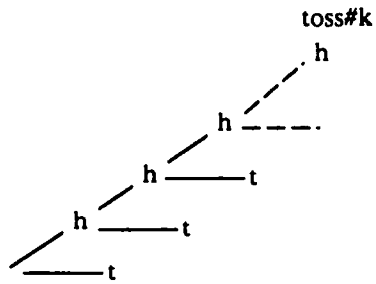


Fig. 4.2

$$\psi(N) = q\psi(1+N) + pq\psi(2+N) + p2q\psi(3-N) ++ p^{k-1}q\psi(k+N) + p^{k}\psi(k) \qquad (4.23)$$

Setting $\psi(N) = z^N$ results in (4.13a).

Shifting each term of (4.23) by 1 and multiplying it by p results in

$$p\psi(1+N) = pq\psi(2+N) + p^{2}q\psi(3+N) + p^{3}q\psi(4+N) ++ p^{k}q\psi(1+k+N) + p^{k+1}\psi(1+k) \quad (4.24)$$

A comparison of (4.23) and (4.24) results in

$$\psi(N) = p\psi(1+N) + q\psi(1+N) + p^{k}q\psi(1+k+N) + p^{k}\psi(k) - p^{k+1}\psi(1+k) \qquad (4.25)$$

Setting $\psi(N) = z^N$ produces (4.13b).

Section 5 Olio

In this section we analyze several examples which are of interest because they show the scope of the omni-method or because of some methodological aspects not previously elaborated.

Example 1 Toss a coin till at least 2 out of 3 consecutive tosses are h; find N.

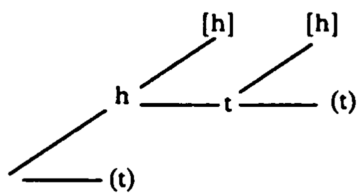


Fig. 5.1

Note that in the path h–t–t in Fig. 5.1 the h is not "operative" because it gets lost upon the very nearest toss. From the tree we read directly

$$\psi(N) = p^2\psi(2) + p^2q\psi(3) + q\psi(1+N) + pq^2\psi(3+N) \tag{5.1}$$

and from (5.1) we have

$$N_j = p^2\delta(j-2) + p^2q\delta(j-3) + qN_{j-1} + pq^2N_{j-3} \tag{5.2}$$

$$N_2 = p^2;\ N_3 = p^2q,\ N_4 = p^2q^2;\ N_5 = p^2q^2;\ N_6 = p^2q^3 + p^3q^3 \text{ etc.,}$$

and

$$\psi(H,T) = p^2\psi(2,0) + p^2q\psi(2,1) + q\psi(H,1+T) + pq^2\psi(1+H,2+T) \tag{5.3}$$

$$N_{ij} = p^2\delta(i-2)\delta(j) + p^2qN_{21} + pq^2N_{i,j-1} + pq^2N_{i-j,j-2}$$

When we toss till at least 2 our of k consecutive tosses show h we have Fig. 5.2.

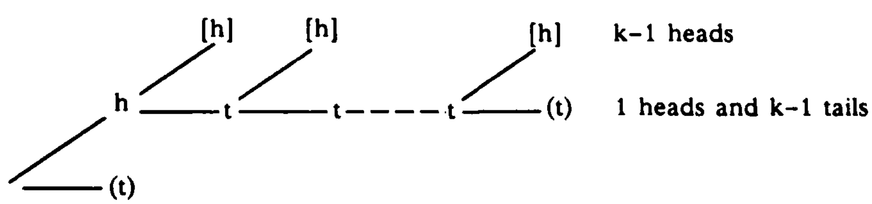


Fig. 5.2

We read off the diagram that

$$\psi(N) = p^2\psi(2) + p^2q\psi(3) ++ p^2q^{k-2}\psi(k) + pq^{k-1}\psi(k+N) + q\psi(1+N) \tag{5.4}$$

The reader should be alerted to the fact that the problem "toss till you get j heads among k successive tosses and count N" is generally more complex combinatorially and cannot be represented by a single finite renewal tree; even if $j = 3$ and $k = 5$ the combinatorics call for the general furkation method or a different approach altogether.

The reader should also verify that the renewal tosses in Fig. 5.1 and in Fig. 5.2 cease to be renewal tosses for the r.v r (number of changes from h to t and from t to h). The general furkation method is needed to analyze the project mentioned.

Example 2 Model {N;ht} Revisited

We have dealt with this model in Section 3 already. Now we derive the recursion for the probabilities (from which we can get the omni-equation) by first finding a few initial probabilities. Suppose that we can justify the following form for that recursion

$$(N)_j = pq\delta(j-2) + c_1(N)_{j-1} + c_2(N)_{j-2} \tag{5.5}$$

We can evaluate c_1 and c_2 if we figure out directly $(N)_3$; we know that $(N)_2 = pq$, the speediest realization of ht. It is easy to see that $(N)_3 = (N)_2$ since the triplets which lead to $N = 3$ are hht and tht; these two jointly have the same probability as ht. From (5.5) we now have

$$(N)_3 = c_1(N)_2 + c_2(N)_1 = c_1(N)_2 \text{ since } (N)_1 = 0$$

But we know that $(N)_3 = (N)_2$ and therefore $c_1 = 1$. And since in (5.5) $pq + c_1 + c_2 = 1$ (this we know to be true since (5.5) is transformable into an omniequation with identical coefficients) we have $c_2 = -pq$. This agrees with (3.7). If we do not know the number of coefficients in (5.5) (these are $c_0 + c_1$ and c_2) we can keep computing them until $c_0 + c_1 ++ c_k = 1$. But even with this procedure it is thinkable that, e.g. $c_{k+1} = 1$. Thus the method is incomplete so far, it is a plausibility.

Let us consider now the model {N;hth}. We assume that the form of the probability recursion is

$$(N)_j = p^2q\delta(j-3) + c_1(N)_{j-1} + c_2(N)_{j-2} + c_3(N)_{j-3} \tag{5.6}$$

a plausible assumption in view of our experience with omni-equations.

We calculate successively (directly from the combinatorics of the model, not from (5.6))

$$(N)_3 = p^2q \; ; \; (N)_4 = (N)_3 \; ; \; (N)_5 = (1-pq)(N)_3 \text{ and } (N)_6 = (1-pq - p^2q)(N)_3 \tag{5.7}$$

From (5.6) and (5.7) we have

$$(N)_4 = c_1(N)_3 \; ; \text{ hence } c_1 = 1$$

From (5.6) and (5.7) we further have

$$(N)_5 = c_1(N)_4 + c_2(N)_3 = c_2(N)_3 = (1-pq)(N)_3 \; ; \text{ therefore } c_2 = -pq$$

Equation (5.6) can now be written

$$(N)_j = p^2q\delta(j-3) + (N)_{j-1} - pq(N)_{j-2} + c_3(N)_{j-3}$$

Since the coefficients on the right-hand side add up to one we have $c_3 = pq^2$; (5.6) is now

$$(N)_j = pq\ (j-3) + (N)_{j-1} - pq(N)_{j-2} + pq^2(N)_{j-3} \tag{5.8}$$

We can compute $(N)_6$ from (5.8) and verify that it agrees with the value given in (5.7).

The omni-equation corresponding to (5.8) is (cf. Table 3.1)

$$\psi(N) = pq\psi(3) + \psi(1-N) - pq\psi(2+N) + pq^2\psi(3+N) \tag{5.9}$$

Equation (5.9) can be transposed into an equation in H,T.

Example 3 Consider the model {N; h,odd #t,h}. The fork equations for this model are

$$\psi(N) = p\psi(1+N_h) + q\psi(1+N), \text{ since } \psi(N_t) = \psi(N) \tag{5.10}$$

$$h \begin{smallmatrix} \nearrow h \\ \longrightarrow t \end{smallmatrix} \quad \psi(N_h) = p\psi(1+N_h) + q\psi(1+N_{h\,t}), \text{ since } \psi(N_{h\,h}) = \psi(N_h) \qquad (5.11)$$

$$ht \begin{smallmatrix} \nearrow h \\ \longrightarrow t \end{smallmatrix} \quad \psi(N_{h\,t}) = p\psi(1) + q\psi(1+N_h), \text{ since } \psi(N_{h\,t\,t}) = \psi(N_h) \qquad (5.12)$$

Only three fork equations suffice for a problem which at first appears combinatorially complex. From the above equations we easily get

$$\bar{N} = \frac{1}{p^2q} + \frac{2}{p}\ ;\ \bar{N}_h = \frac{1}{pq} + \frac{1}{p}\ ;\ \bar{N}_{ht} = \frac{2}{p} \qquad (5.13)$$

For $p = q = 1/2$ (5.13) becomes $\bar{N} = 8$; $\bar{N}_h = 6$; $\bar{N}_{h\,t} = 4$.

From (5.10), (5.11) and (5.12) we can derive all higher moments and the probabilities for N, for N_h for N_t and for $N_{t\,h}$. We can also derive the gf's for N, for N_h, for N_t and for $N_{h\,t}$ by setting in those equations $\psi(N) = z^N$, etc. We thus get

$$Ez^N = pzEz^{N_h} + qzEz^N$$

$$Ez^{N_h} = pzEz^{N_h} + qzEz^{N_{ht}}$$

$$Ez^{N_{ht}} = pz + qzEz^{N_h}$$

from which we get

$$Ez^N = \frac{p^2qz^3}{1 - z+q(p-q)z^2 + q^3z^3} \qquad (5.14)$$

$$Ez^{N_h} = \frac{pqz^2}{1 - pz - q^2z^2} \qquad (5.15)$$

$$Ez^{N_{ht}} = \frac{pz(1-pz)}{1 - pz - q^2z^2} \qquad (5.16)$$

The last three equations can be easily transposed into omni-equations (c.f. Table 3.1)

$$\psi(N) = \psi(1+N) - q(p-q)\psi(2+N) - q^3\psi(3+N) + p^2q\psi(3) \qquad (5.14a)$$

$$\psi(N_h) = \psi(1+N_h) + p\psi(1+N_h) + q^2\psi(2+N_h) + pq\psi(2) \tag{5.15a}$$

$$\psi(N_{h\,t}) = p\psi(1+N_{h\,t}) + q^2\psi(2+N_{h\,t}) + p\psi(1) - p^2\psi(2) \tag{5.16a}$$

The three above equations, or the equations (5.14), (5.15) and (5.16) can be transposed into the corresponding equations for the probabilities. The reader interested in converting (5.14) or (5.14a) into an equation for (H,T) should make it upright. A neat exercise is also to derive (5.15) by assuming (and plausibly justifying) that

$$\psi(N) = \alpha_0\psi(3) + \alpha_1\psi(1+N) + \alpha_2\psi(2+N) + \alpha_3\psi(3+N) \tag{5.17}$$

and deriving the α_i from several initial values (reasoned out independently) $(N)_j$. (Transposing (5.17) into a probability equation may ease the exercise.)

Example 4 Model {N;h,k*t,h}

In this model k is a fixed positive integer. It is not difficult to reason out that

$$(N)_{k+2} = p^2q^k = (N)_{k+3} = = (N)_{2k+2} \tag{5.18}$$

We assume that

$$(N)_j = p^2q^k\delta(j-k-2) + \alpha_1(N)_{j-1} + \alpha_2(N)_{j-2} ++ \alpha_{k+2}(N)_{j-k-2} \tag{5.19}$$

Thus, for k = 1 we have $(N)_3 = p^2q = (N)_4$ and $(N)_j = p^2q\delta(j-3) + \alpha_1(N)_{j-1} + \alpha_2(N)_{j-2} + \alpha_3(N)_{j-3}$. This equation is plausible since there are, as is easy to check, three fork equations when k=1.

It is further easy to see that $\alpha_1 = 1$ and $\alpha_2 = = \alpha_k = 0$ when $k \geq 2$ by comparing the $(N)_j$ from (5.18) with the $(N)_j$ computed from (5.19) for j = k+2, k+3, ..., 2k+2. We further find that

$$(N)_{2k+3} = (1-pq^k)(N)_{k+2} \quad \text{and} \quad (N)_{2k+4} = (1-pq^k-p^2qk)(N)_{k+2} \tag{5.20}$$

We get eventually

$$(N)_j = p^2q^k\delta(j-k-2) + (N)_{j-1} - pq^k(N)_{j-k-1} + pq^{k+1}(N)_{j-k-2} \tag{5.21}$$

which specifies the α_i in (5.19). We note that the coefficients on the righthand side sum to 1, as they should. The requirement that the coefficients on the left-hand side and right-hand side sum to the same value follows from the term-by-term correspondence with the omni-equation

$$\psi(N)=p^2q^k\psi(k+2) +\psi(1+N)) - pq^k\psi(k+1+N) + pq^{k+1}\psi(k+2+N) \tag{5.22}$$

by setting $\psi(\cdot) = 1$.

Example 5 Model {N; k*h, l*t} We state the following equation for the model {N;k*h,l*t} without derivation:

$$(N)_j = p^k q^l\, \delta(j-k-l) + (N)_{j-1} - p^k q^l\, (N)_{j-k-1} \tag{5.23}$$

from which follow (cf Table 3.1)

$$\psi(N) = p^k q^l\, \psi(k+l) + \psi(1+N) - p^k q^l\, \psi(k+l+N) \tag{5.24}$$

and

$$Ez^N = \frac{p^k q^l}{1 - z + p^k q^l\, z^{k+l}} \tag{5.25}$$

When $k=l=1$ then (5.23) is the same as (3.7) and our model is the familiar {N;ht}.

Note: Consider a string σ of h's and t's in some order. Assume that in this string no toss, except for the last one, can be a stopping toss whatever sequence of h's and t's would precede σ. Then (5.23), (5.24) and (5.25) are valid for {N; σ} with k=# heads in the string σ and l=# tails in σ.

Example 6 Target {hh or tt}

Feller, p. 327, (8.2) derives a gf for the number of tosses, N, to reach a string of r heads (event ϵ_1) or a string of ρ tails (event ϵ_2) whichever comes first. He also, (8.6), derives gf's for the probabilities of "reaching ϵ_1 in n tosses without hitting ϵ_2 first" and of "reaching ϵ_2 in n tosses without hitting ϵ_1 first."

We develop a set of fork-equations for the random vector (N,a) where a = 1, or else 2, if the tossing results in event ϵ_1, or else in, event ϵ_2; the random variables a_h and a_t are conditioned upon having just tossed h or else t. Our example is modest in that we select tiny target strings but the method is indicated clearly. We add that the forking method is applicable to more than two alternative events which, moreover, may be strings other than heads only or tails only. We can choose, e.g. ϵ_1 = hhtth, ϵ_2 = tthtt and ϵ_3 = hhhttt; and instead, or in addition to, N we can count the number of doublets ht.

For the model {N;hh or tt} the fork-equations are

$$\begin{array}{l} \nearrow h \\ \rightarrow t \end{array} \qquad \psi(N,a) = p\psi(1+N_h,a_h) + q\psi(1+N_t,a_t) \tag{5.26}$$

$$h \begin{array}{l} \nearrow [h] \\ \rightarrow t \end{array} \qquad \psi(N_h,a_h) = p\psi(1,1) + q\psi(1+N_t,a_t) \tag{5.27}$$

$$t \begin{array}{l} \nearrow h \\ \rightarrow [t] \end{array} \qquad \psi(N_t,a_t) = p\psi(1+N_h,a_h) + q\psi(1,2) \tag{5.28}$$

Thus, along with (N,a) we analyze (N_h,a_h) and (N_t,a_t); this is in the nature of the furcation method. From (5.26), (5.27) and (5.28) we can derive the distribution and the moments of interest. We can with ease transpose these equations into equations in (H,T,a), (H_h,T_h,a_h) and (H_t,T_t,a_t). We can also derive, with some fair amount of algebra, the uncoupled equations (5.29), (5.30) and (5.31):

$$\begin{aligned} \psi(N,a) &= pq\psi(2+N,a) + p^2\psi(2,1) + pq^2\psi(3,2) \\ &\quad + qp^2\psi(3,1) + q^2\psi(2,2) \end{aligned} \tag{5.29}$$

$$\psi(N_h a_h) = pq\psi(2+N_h,a_h) + p\psi(1,1) + q^2\psi(2,2) \tag{5.30}$$

$$\psi(N_t,a_t) = pq\psi(2+N_t,a_t) + p^2\psi(2,1) + q\psi(1,2) \tag{5.31}$$

From (5.29), which is upright with respect to the N-argument, we get easily (5.32):

$$\begin{aligned} \psi(H,T,a) &= pq\psi(1+H,1+T,a) + p^2\psi(2,0,1) \\ &\quad + pq^2\psi(1,2,2) + qp^2\psi(2,1,1) + q^2\psi(0,2,2) \end{aligned} \tag{5.32}$$

We can do the same for (5.30) and (5.31), and we can likewise transpose (5.22), (5.23) and (5.24) each into a joint trivariate probability distribution. The interested reader should have no difficulties doing it.

From (5.29) we find

$$\psi(a) = \frac{p^2(1+q)}{1-pq}\ \psi(1) + \frac{q^2(1+p)}{1-pq}\ \psi(2) \tag{5.33}$$

and hence

$$\Pr(a=1) = \frac{p^2(1+q)}{1-pq} \quad \text{and} \quad \Pr(a=2) = \frac{q^2(1+p)}{1-pq} \tag{5.34}$$

Bibliography

Botta, R. F. and C. M. Harris, Approximation with Generalized Hyperexponential Distributions: Weak Convergence Results, Queueing Systems 2 (1986).

Feller, William, An Introduction to Probability Theory and its Applications, vol. 1, Third Edition, 1968.

Gross and Harris, Fundamentals of Queueing Theory, 2nd edition, Wiley, 1985.

Krakowski, Martin, System Size in Some Variants of M/G/1, Report No. GMU/49146/104, July 1986

Krakowski, Martin, M/G/1 Subject to an Initial Quorum of Customers, Report No. GMU/49146/105

DISTRIBUTION LIST

Copy No.	
1	Office of Naval Research 800 North Quincy Street Arlington, VA 22217 Attention: Scientific Officer, Statistics and Probability Mathematical Sciences Division
2	ONR Resident Representative Joseph Henry Building, Room 623 2100 Pennsylvania Avenue, N.W. Washington, DC 20037
3 - 8	Director, Naval Research Laboratory Washington, DC 20375 Attention: Code 2627
9 - 20	Defense Technical Information Center Building 5, Cameron Station Alexandria, VA 22314
21 - 29	C. M. Harris
30	GMU Office of Research